

Article

Power Grid Violation Action Recognition via Few-Shot Adaptive Network

Lingwen Meng ¹, Lan Zhang ^{2,*} , Guobang Ban ¹, Shasha Luo ¹ and Jiangang Liu ¹

¹ Electric Power Research Institute of Guizhou Power Grid Co., Ltd., Guiyang 550002, China; menglw@gzsy.csg.cn (L.M.); banforks@163.com (G.B.); luoshasha931@163.com (S.L.); 15985305614@163.com (J.L.)

² School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

* Correspondence: zl2022@tju.edu.cn

Abstract: To address the performance degradation of violation action recognition models due to changing operational scenes in power grid operations, this paper proposes a Few-shot Adaptive Network (FSA-Net). The method incorporates few-shot learning into the network design by adding a parameter mapping layer to the classification network and developing a task-adaptive module to adjust the network parameters for changing scenes. A task-specific linear classifier is added after the backbone, allowing the adaptive generation of classifier weights based on the changing task scene to enhance the model's generalizability. Additionally, the model uses a strategy of freezing the backbone network and iteratively updating only certain module parameters during training in order to minimize training costs. This approach addresses the challenge of iteratively updating difficulties in the original model, which are caused by limited image data following scene changes. In this paper, 2000 samples under power grid scenarios are used as the experimental dataset; the average recognition accuracy for violation actions is 81.77% for images after scene changes, which represents a 4.58% improvement when compared to the ResNet-50 classification network. Furthermore, the model's training efficiency is enhanced by 40%. The experimental results show that the method enhances the performance of the violation action recognition model before and after scene changes and improves the efficiency of the iterative model by updating with a smaller sample size, lower model design cost, and lower training cost.



Academic Editor: Dorin Petreus

Received: 8 November 2024

Revised: 11 December 2024

Accepted: 13 December 2024

Published: 30 December 2024

Citation: Meng, L.; Zhang, L.; Ban, G.; Luo, S.; Liu, J. Power Grid Violation Action Recognition via Few-Shot Adaptive Network. *Electronics* **2025**, *14*, 112. <https://doi.org/10.3390/electronics14010112>

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: power grid violation; few-shot learning; action recognition; parameter mapping layer; task adaptive module; task-specific linear classifier

1. Introduction

With the expansion of power grid construction and renovation projects, electric power construction sites have become widely distributed and often involve multiple construction teams, making safety risk control increasingly challenging. Violations in grid operation scenarios typically stem from operator carelessness, non-compliance, unfamiliarity with procedures, or equipment failures [1,2]. Such behaviors can have significant impacts on the safety and reliability of a power system [3–5]. Non-compliance with safety procedures in a grid scenario can result in severe consequences, including substantial property damage and risks to personal safety. For instance, improper use of grid equipment or failure to conduct regular inspections can cause hazardous conditions such as partial discharges [6]. Adherence to safety procedures is crucial for ensuring the continuous and stable operation of a power grid. Traditional identification and prevention of grid operation violations

currently require significant human resources [7,8]. Additionally, many potential violations are challenging to detect. Consequently, there is an urgent need for intelligent monitoring and control systems in grid operation behavior identification to enhance the security, compliance, and efficiency of power system operations and maintenance [9–11]. With the advancement of machine learning and deep learning, intelligent violation action recognition technology has increasingly been adopted in the power grid sector. However, the operating environment often undergoes subtle changes in real grid scenarios. For instance, the operation scene is frequently influenced by complex environmental conditions, including adverse weather, low light, dust, and vegetation [12]. These objective factors can cause slight variations in the operation scene, resulting in the degradation of the original violation recognition model's performance in grid operation scenarios. Additionally, the limited amount of image data following a scene change complicates the iterative updating of the original model, hindering its adaptability to the modified operation scene. Therefore, it is essential to develop a model with robust generalization and rapid iterative updating capabilities.

Currently, there are fewer studies on few-shot learning in grid scenarios [13,14]. Furthermore, these methods are designed for processing one-dimensional data, such as in predicting future electricity loads based on historical load data. These models cannot process two-dimensional image inputs and are, therefore, unsuitable for violation action recognition in grid scenarios. In summary, the industry currently lacks few-shot models tailored for violation action recognition in grid scenarios. Our approach provides an idea for few-shot learning in grid scenarios. The main contributions of this paper are as follows:

To address the shortcomings in the existing technology, this paper proposes a few-shot adaptive network for grid violation action recognition. We add a parameter mapping layer to the convolutional block of the classification network, which maps the features to adapt to the task in the changing scenarios. Additionally, we integrate a task adaptation module parallel to the backbone network, which supplies weights to the parameter mapping layer based on the specific task scenario. To further enhance the model's generalization, we develop task-specific linear classifiers that enable the model to generate distinct classifier weights for different scenarios. Finally, we employ a training strategy that involves freezing the backbone network and fine-tuning selected structures to lower iterative update costs. Extensive experiments have demonstrated that our approach increases the accuracy and efficiency of potential violation recognition in typical power grid operation scenarios.

2. Related Works

2.1. Grid Violation Action Recognition

Currently, the primary methods for safety risk monitoring in field operations are manual safety monitoring and intelligent monitoring [15]. The currently employed manual safety monitoring methods primarily involve assigning dedicated supervisors to oversee operators' behavior and activities [16]. However, supervisors cannot guarantee comprehensive supervision of operators and are equally susceptible to external factors that may distract them, potentially resulting in safety incidents. With advances in computer technology, some researchers have utilized image processing methods for safety risk identification. Cai et al. [17] proposed an image recognition method for substation signage using traditional image processing techniques, which helps prevent substation operators from entering incorrect compartments, thereby ensuring their safety. Long et al. [18] introduced a helmet detection method based on a deep convolutional neural network (DCNN), capable of detecting instances of entering the workplace without a helmet using monitoring data, thereby enhancing operator safety. Liu et al. [19] developed a universal pointer meter detection and identification method using a target detection model and FAST R-CNN

neural network, enabling automatic meter reading and minimizing operator exposure to high-voltage environments. Wang et al. [20] introduced a new framework for substation video monitoring using artificial intelligence technology to reduce human accidents and provide functionalities such as helmet detection, automatic fire alarms, and automated alerts for staff entering hazardous areas. However, all of these violation recognition algorithms are trained using large-scale datasets. These datasets are typically collected in fixed grid scenarios, which means the models are primarily effective in a single scenario, leading to significant performance degradation when confronted with changing grid environments. Moreover, new grid scenarios with limited samples often hinder the model's ability to adapt quickly to the new conditions. Therefore, there is an urgent need for an effective method that enables models to be updated quickly and iteratively, even with limited samples.

2.2. Few-Shot Learning

Few-shot learning refers to scenarios where the number of samples is limited, restricting the model's training to this constrained data. Current research on few-shot learning primarily investigates three dimensions: data, models, and optimization algorithms. Data-centric approaches enhance learning by augmenting samples and feature information. Model-centric strategies aim to reduce the hypothesis space through structural and parameter design. Optimization-oriented methods increase the likelihood of identifying the optimal hypothesis by modifying search strategies within a defined hypothesis space. Limited data volumes primarily affect the accuracy and stability of feature selection throughout the process. Data augmentation diversifies datasets without incurring additional sampling costs, thereby preventing overfitting and enhancing the utility of small datasets [21]. Generative Adversarial Networks (GANs) concurrently train two adversarial models: a generator network that creates artificial data by capturing the original data distribution from noise and a discriminator network that learns to differentiate between generated and real data [22]. However, standard GANs face challenges such as training instability, mode collapse, and difficulties in evaluation [23,24], which hinder the generator's ability to learn diverse data distributions [24]. Typical optimization-oriented small sample learning methods, such as Model-Agnostic Meta-Learning (MAML) [25], enable quick adaptation to new tasks through cross-task training. Despite its successes, MAML requires computationally expensive second-order derivatives for updates. In this paper, we tackle the small sample problem in grid scenarios from the perspective of model design.

3. Methods

3.1. Few-Shot Adaptive Network

With the wide application of deep learning techniques in the field of computer vision, significant progress has been made in action recognition techniques. However, when these action recognition models are applied in real grid scenarios, they often encounter challenges such as lighting changes, adverse weather conditions, and vegetation occlusion; consequently, adapting these deep models to new scenarios frequently results in a significant decrease in action recognition accuracy. Additionally, even if some samples are collected for the new environment post-change, issues related to model iteration and updating, such as insufficient sample size and excessive training costs, may arise. In light of these challenges present in real grid scenarios, this chapter proposes few-shot adaptive networks. By reducing sample size, model design costs, and training costs, this approach enhances the performance of the behavior recognition model and further improves the efficiency of iterative model updates.

Firstly, the feature extraction component of the network is outlined, including the original classification network structure, the newly designed parameter mapping layer,

and the task adaptive module. Subsequently, the classifier component of the network is described, primarily consisting of task-specific linear classifiers. Finally, the training strategy of the model will be described, detailing which parameters will be trained.

3.1.1. Feature Extraction Backbone

As shown in Figure 1, the feature extraction backbone network of the model comprises a modified ResNet backbone and four parallel task-adaptive modules.

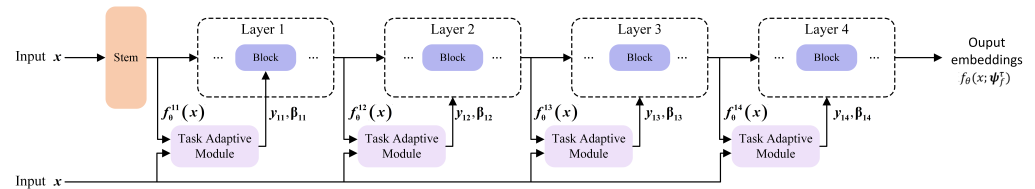


Figure 1. The framework of the feature extraction backbone.

The ResNet backbone selected is ResNet-50 [26], with the addition of a parameter mapping layer. The parameter mapping layer adaptively maps features extracted from the convolutional block, ensuring they are more suited to the current task scenario. To minimize model complexity, the parameter mapping layer is included only after the final BN layer in the first block of each layer. Each parameter mapping layer, as shown in Figure 2, has two parameters: γ_{1i} and β_{1i} . The calculation formula is as follows:

$$F_i(f_i; \gamma_{1i}, \beta_{1i}) = \gamma_{1i}f_i + \beta_{1i}, \quad (1)$$

where f_i represents the unmapped features output from Block 1 in the i th layer, F_i represents the mapped features output from Block 1 in the i th layer, and γ_{1i} and β_{1i} denote the parameters of the parameter mapping layer in Block 1 of the i th layer. The feature extraction backbone of ResNet-50 consists of 4 layers, necessitating the addition of 4 parameter mapping layers. The parameter mapping layer adaptively refines the features extracted by the module. When encountering images from altered scenes, although the features extracted by the original module may no longer be suitable for accurate classification, the parameter mapping layer adjusts these features to align with the new scene.

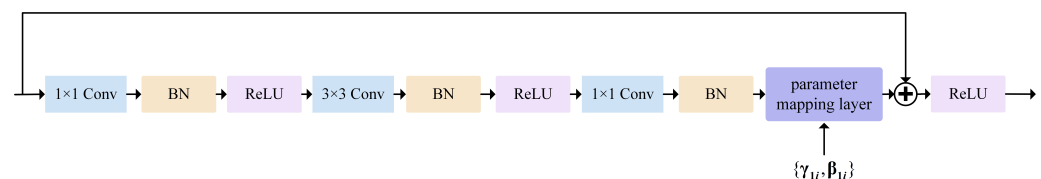


Figure 2. The structure of improved ResNet block.

The two parameters of each parameter mapping layer are generated by the parallel task adaptive module. The structure of the task adaptive module is shown in Figure 3. The task adaptive module consists of a stack of 3×3 convolutional layers, a max-pooling layer, an average pooling layer, a channel concatenation layer, a linear layer, and a ReLU activation function. Here, convolution refers to the combined operation of convolution, batch normalization, and the ReLU activation function. The task adaptive module takes two inputs: the input feature $f_{\theta}^{1i}(x)$ from Block 1 in the i th layer of ResNet, and the image x from the slightly changed scene. The image x from the slightly changed scene first passes through the convolutional layer and max-pooling layer for initial feature extraction. It is then concatenated with the feature $f_{\theta}^{1i}(x)$ from ResNet, and finally, the parameters γ_{1i} and β_{1i} required for the parameter mapping layer are output through the fully connected layer and the ReLU layer. Under this structural design, the parameters of the parameter mapping

layer are derived from a combination of features from the actual image and those from the original scene. In contrast to directly defining learnable parameters in the parameter mapping layer, the parameters generated in the task adaptation module are more effective and interpretable.

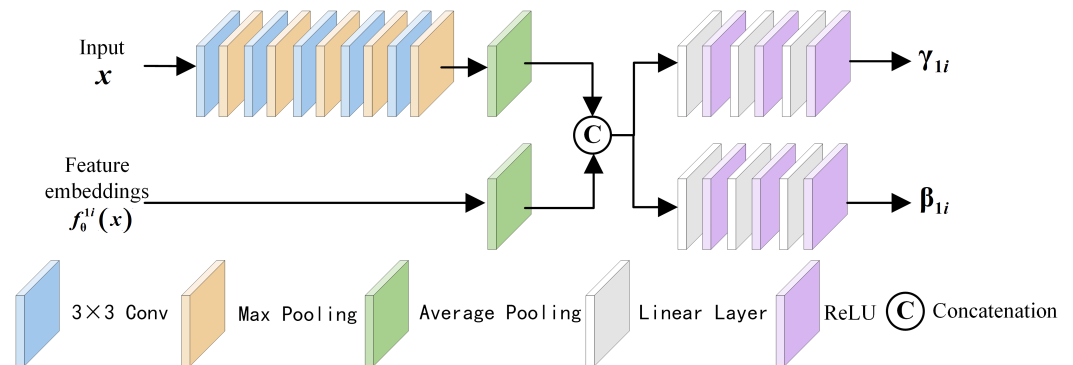


Figure 3. The architecture of the task adaptive module.

3.1.2. Task-Specific Linear Classifier

In traditional classification networks, the classifiers are often composed of simple, fully connected layers. In this subsection, the task-specific linear classifier is designed to generate the weight matrices of the classifiers using a specific weight generation method. The structure is shown in Figure 4.

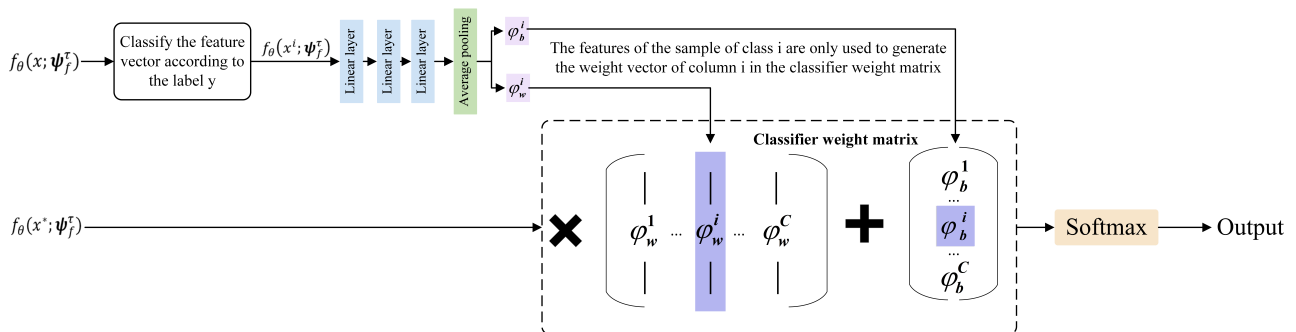


Figure 4. The architecture of the task-specific linear classifier.

Specifically, the task-specific linear classifier consists of a weight matrix. Specifically, the task-specific linear classifier consists of a weight matrix and a Softmax function. The weight matrix consists of $2C$ parameters φ_w^i and φ_b^i . During the training process, few-shot images are fed into the feature extraction backbone, and task-specific features $f_\theta(x; \psi_f^\tau)$ are extracted, where θ represents the original parameters of ResNet-50, x represents the few-shot images, and ψ_f^τ represents the parameters of the new structure added for the specific task τ . The extracted features are classified according to the image labels i . The features $f_\theta(x^i; \psi_f^\tau)$ of the i th class of few-shot images are used to generate φ_w^i and φ_b^i in the i th column of the weight matrix of the linear classifier after passing through three fully connected and pooling layers. There are eight columns in the weight matrix corresponding to eight types of violations. When the images in the training batch belong to category i , only the parameters in column i of the weight matrix are iteratively updated, while the parameters in other columns remain frozen. The task-specific classifier uses this strategy of updating corresponding class weights to maintain high iterative update efficiency when dealing with a small sample size.

Finally, in the testing phase, a new sample x^* is input, and task-specific features $f_\theta(x; \psi_f^\tau)$ are extracted through the feature extraction backbone. The features are then

multiplied by the classifier's weight matrix to yield a vector of size 8×1 , which is passed through Softmax to obtain the probability of the sample belonging to each category.

3.2. Network Training Optimization Strategies

In this subsection, the main focus will be on the model's training strategy and parameter update strategy. The training objective function defines the model's learning goal, typically achieved by minimizing the loss function. Network model optimization involves using an algorithm to adjust the model's parameters to minimize the loss function. Through iterative optimization, the model progressively learns data features, enhancing its performance.

The ResNet-50 classification network is initially trained with sufficient sample data, serving two purposes: first, to establish a baseline model for evaluating and comparing; second, to provide initial weighting parameters for the proposed few-shot adaptive network.

Subsequently, the classifier portion of the ResNet-50 classification network is removed, retaining the feature extraction backbone, and the designed parameter mapping layer, task adaptive module, and task-specific linear classifier are incorporated. The model is then trained using few-shot images after slight scene changes, with the parameters of the original ResNet-50 feature extraction backbone frozen, while only the parameters in the newly added parameter mapping layer, task adaptive module, and task-specific linear classifier are iteratively updated.

4. Experiment

In order to validate the effectiveness and feasibility of the proposed method in this paper, the proposed method is trained and tested on a dataset with 2000 labeled samples, and comparative experiments with the baseline algorithm are carried out. These validation results demonstrate that the proposed method provides a solid foundation for further research and practical applications.

4.1. Violation Action Dataset Construction

In a typical grid operation scenario, representative sample data on behaviors such as working at heights, checking electrical equipment, and hanging earth wires under varying conditions (e.g., changes in lighting, time of day, and weather) were collected and labeled by professionals to create a dataset of 2000 samples. The labeled violation action categories K were defined as eight types: smoking, not wearing a safety helmet, not wearing work clothes, not wearing a safety harness, not wearing insulated gloves, not wearing insulated shoes, sitting or crossing a railing at the edge of a high platform or hole, and throwing tools or materials during high-altitude operations. Each sample includes at least one violation action. The dataset comprises 2000 images capturing workers performing tasks at various times of the day (day, dusk, night) and in different work areas, including distribution rooms and substations. These images comprehensively represent the common areas and work sites in actual power grid scenarios. In addition, the scale of individuals in the images varies significantly, reflecting the varying distances and angles of surveillance cameras in real-world scenarios. In this dataset, 200 images with poor lighting conditions were selected to represent a grid scenario with slight changes. Before iterative training, datasets with varying lighting conditions were randomly divided into training and validation sets at an 8:2 ratio. Specifically, 1440 images were used for training and 360 for validation before scene changes, while 160 images were designated for training and 40 for testing after the scene change. Several images of the dataset are shown in Figure 5.

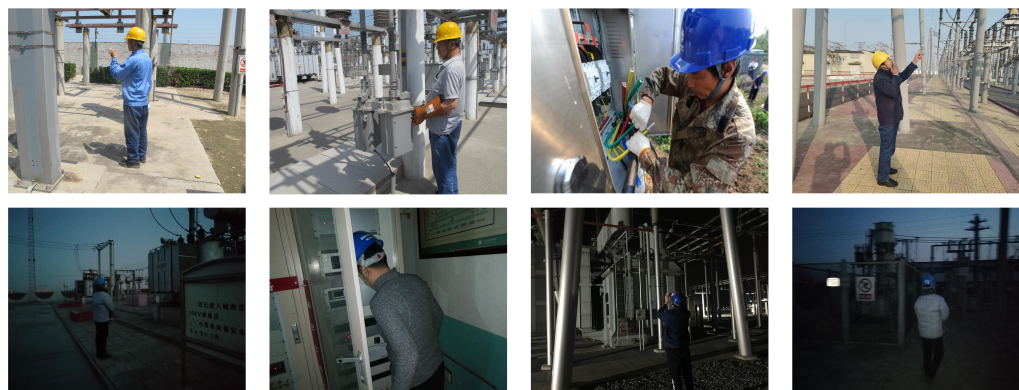


Figure 5. Several images of random selection in the dataset.

4.2. Implementation Details and Evaluation Metrics

The software configuration of the experimental system included Python 3.8, PyTorch 1.9.0, CUDA 11.2, and PyCharm 2023.2, while the hardware configuration comprised an RTX 3090 graphics card (24 GB VRAM). During the training phase, all training samples were scaled to a resolution of 224×224 before being input into the network. These samples underwent various data augmentation techniques, including random cropping and horizontal flipping. Random cropping involved selecting random rectangular regions from the images, retaining 80% to 100% of the original area, followed by resizing to the original dimensions to simulate diverse visual contexts. Horizontal flipping mirrored the images to enhance invariance to orientation. The SGD optimizer was employed for training the ResNet-50 classification network, with an initial learning rate set to 1×10^{-3} and a cosine learning rate decay strategy. The batch size (N) was set to 32, and training was conducted over 50 epochs. When training on small sample images after slight scene changes, the parameters of the feature extraction backbone of the original ResNet-50 were frozen, and only the parameters in the newly added parameter mapping layer, task adaptive module, and task-specific linear classifiers were updated. The SGD optimizer was still used for this phase, with an initial learning rate of 5×10^{-4} and a batch size (N) of 16. After the 10th and 20th epochs, the learning rate was reduced to 0.1 times its original value, and training continued for a total of 30 epochs. This parameter selection is driven by the limited number of samples in the new scenario, where the model requires fine-tuning rather than full retraining; hence, both the learning rate and the number of iterations are set to relatively low values.

In the experiments, Accuracy is used to evaluate the action recognition results, and it is defined as the proportion of correctly classified samples, which is calculated as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \quad (2)$$

where TP represents the number of true positive samples, TN represents the number of true negative samples, FP represents the number of false positive samples, and FN represents the number of false negative samples.

4.3. Results and Analysis

4.3.1. Ablation Study

In this chapter, ablation experiments were conducted, as shown in Table 1, where the designed components were gradually added under the baseline of ResNet-50 in order to analyze the role of the components in the methodology of this chapter. Since the Task Adaptive Module needs to be used in conjunction with the Parameter Mapping Layer, the Task Adaptive Module is used as a proxy for the combination of the two components in

the ablation experiments. It can be seen that when using only the task adaptive module and the task-specific linear classifier, the average accuracy of action recognition is reduced compared to the full model. The best performance is exhibited when all components are used. Therefore, it can be seen that both the task-adaptive module and the task-specific linear classifier provide performance gains.

Table 1. Accuracy comparison of the model on the test set with different components.

Components				Before the Scene Change (%)	After the Scene Change (%)
ResNet	PML	TAM	TSLC		
✓				83.70	77.19
✓	✓			83.90	78.13
✓	✓	✓		84.11	79.84
✓			✓	84.05	78.22
✓	✓	✓	✓	84.49	81.77

4.3.2. Qualitative and Quantitative Assessment Comparison

The proposed method was evaluated separately using test sets from before and after minor scenario changes. Table 2 presents the recognition accuracies of the proposed methods for each category and includes the baseline ResNet-50 for comparison. To ensure a fair comparison, we initialized the ResNet-50 model with weights obtained from training on the original scene and fine-tuned it using images with slight scene variations with a learning rate consistent with the proposed FSA-Net until convergence.

Table 2. Accuracy comparison of the model on the test set for different categories of violation action.

Violation Action	Before the Scene Change (%)		After the Scene Change (%)	
	ResNet	FSA-Net	ResNet	FSA-Net
Smoking	91.13	91.73	85.43	90.21
No safety helmet	91.96	92.68	86.78	91.02
No work clothes	78.56	79.24	72.44	76.38
No safety harness	73.43	73.78	68.57	71.20
No insulated gloves	85.52	85.93	78.32	84.37
No insulated shoes	71.13	72.27	65.24	69.65
Leaning on or over railings	86.36	87.52	80.25	86.02
Throwing implements or materials	91.48	92.78	80.48	85.27

As can be seen from Table 2, the network proposed in this paper has been evaluated on a typical grid operation dataset. The accuracy for the original scene images, which represent the images under adequate lighting conditions, is as follows: smoking 91.73%, not wearing a safety helmet 92.68%, not wearing work clothes 79.24%, not wearing a safety harness 73.78%, not wearing insulated gloves 85.93%, not wearing insulated shoes 72.27%, sitting or crossing a railing at the edge of high platforms or holes 87.52%, and throwing tools or materials during high-altitude operations 92.78%. These results show an improvement of 0.60%, 0.72%, 0.68%, 0.35%, 0.41%, 1.14%, 1.16%, and 1.3%, respectively, compared to the baseline method. For images after slight scene modifications, i.e., images with reduced light, the accuracy rates are as follows: smoking 90.21%, not wearing a safety helmet 91.02%, not wearing work clothes 76.38%, not wearing a safety harness 71.20%, not wearing insulated gloves 84.37%, not wearing insulated shoes 69.65%, sitting or crossing a railing at the edge of high platforms or holes 86.02%, and throwing tools or materials

during high-altitude operations 85.27%, representing improvements of 4.78%, 4.24%, 3.94%, 2.63%, 6.05%, 4.41%, 5.77%, and 4.79%, respectively, over the baseline method. The baseline method used the ResNet-50 classification network, which was first trained on the original scene images, fine-tuned with the images after slight scene changes, and then tested. The results demonstrate that the proposed model not only slightly improves the accuracy of behavior recognition in the original scene but also significantly enhances the accuracy for scenes with slight changes.

To better illustrate the effectiveness of the few-shot adaptive network proposed in this paper for recognizing violation actions during grid operations, this chapter presents qualitative results for some images from the test set, as shown in Figure 6. The three sub-figures in the first row of the figure depict the recognition outcomes for violation actions in the original grid scene, while the three sub-figures in the second row display the recognition outcomes after slight scene modifications (e.g., dark lighting). These results demonstrate that the proposed network effectively recognizes violation actions in grid operations under both original and changed scene conditions.

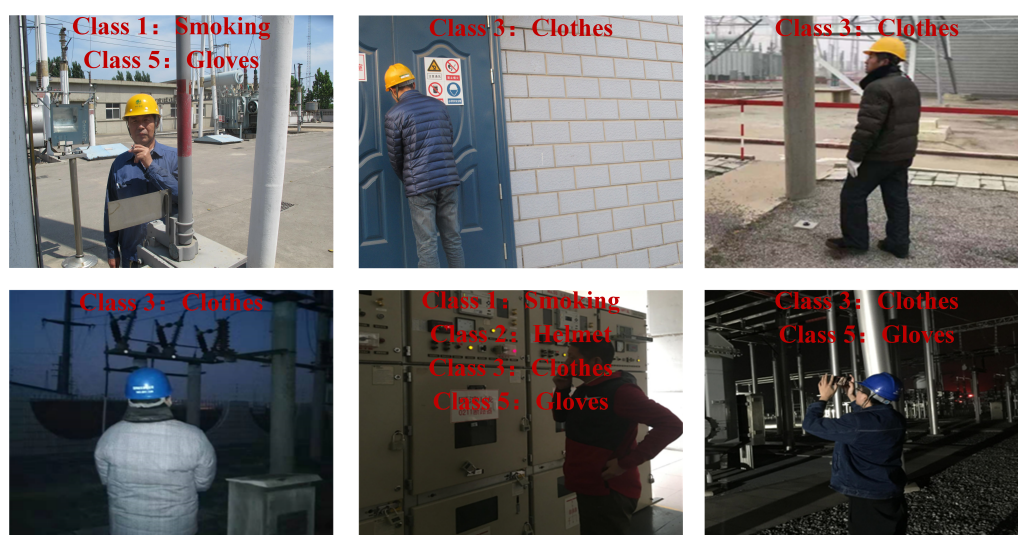


Figure 6. Visualization results of violation action recognition for worker in power grid operation. The three sub-figures in the first row of the figure depict the recognition outcomes for violation actions in the original grid scene, while the three sub-figures in the second row display the recognition outcomes after slight scene modifications (e.g., dark lighting).

4.3.3. Computational Complexity and Efficiency Analyses

This chapter evaluates the complexity of both the baseline and the complete ResNet-50 networks, with the results presented in Table 3. The findings indicate that, compared to the baseline ResNet-50 model, the model proposed in this paper only increases the number of parameters by 1.9 M and the number of operations by 6.06 GFlops. An analysis of these results indicates that the proposed few-shot adaptive network enhances violation recognition accuracy with minimal additional design and training costs.

Table 3. Comparison of parameter number and computation amount of the model.

Model	Params (M)	Flops (G)
ResNet-50	5.34	23.45
FSA-Net	7.24	29.51

To further assess the model's efficiency, we illustrate the number of epochs required for the proposed FSA-Net and the baseline ResNet-50 to train to convergence after scenario

changes. As shown in Figure 7, the results indicate that the proposed FSA-Net converges at 30 epochs, while ResNet-50 requires 50 epochs to reach convergence. This demonstrates that the proposed model is more efficient in fine-tuning during training.

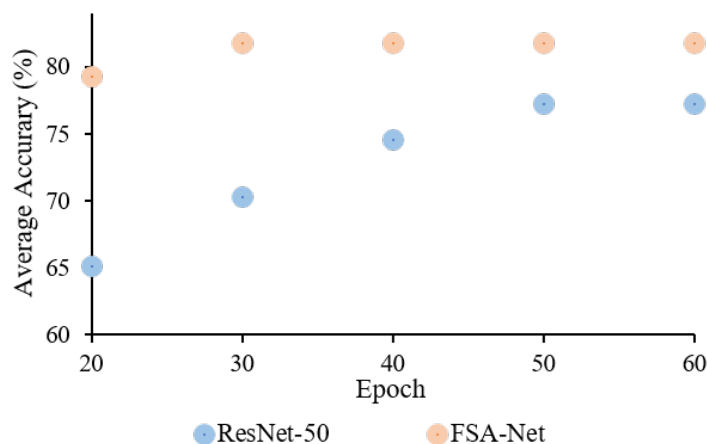


Figure 7. Comparison of the results on different training epochs.

5. Conclusions

This paper proposes a few-shot adaptive network to address the challenges posed by slight variations in operational scenes during grid operations, which result in performance degradation of the original violation action recognition model and limited image data post-scene change, complicating iterative model updates. In response to slight variations in grid scenes, the task adaptive module and parameter mapping layer are developed to make minor adjustments to network parameters specific to the scene and to predict task-specific linear classifier weights. This approach enhances the performance of the behavior recognition model with a smaller sample size, lower model design costs, and reduced training expenses while also improving the efficiency of iterative model updates. On the constructed grid operation dataset, the average recognition accuracy for violation actions is 81.77% for images after scene changes, which represents a 4.58% improvement when compared to the ResNet-50 classification network. Furthermore, the model's training efficiency is enhanced by 40%. Consequently, the model achieves greater accuracy and efficiency in recognizing potential violation behaviors in typical grid operation scenarios. A limitation of this model is that it was tested only under slight variations in lighting conditions. Minor changes in operational scenarios resulting from other conditions will also follow the proposed training strategy. Future work will explore the model's feasibility when scenarios change due to different conditions.

Author Contributions: Conceptualization, L.M. and L.Z.; methodology, L.M., L.Z., G.B., S.L. and J.L.; software, L.Z.; validation, L.Z., G.B., and S.L.; formal analysis, L.Z.; investigation, J.L.; resources, G.B.; data curation, S.L.; writing—original draft preparation, L.Z.; writing—review and editing, L.M., G.B., S.L. and J.L.; visualization, J.L.; supervision, L.M.; project administration, L.M.; funding acquisition, L.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Guizhou Power Grid Co., Ltd., grant number GZKJXM20222320.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are unavailable due to privacy or ethical restrictions.

Acknowledgments: We are grateful for the administrative and technical support of Guizhou Power Grid Co., Ltd.

Conflicts of Interest: Author Lingwen Meng, Guobang Ban, Shasha Luo and Jiangang Liu are employed by the company Electric Power Research Institute of Guizhou Power Grid Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Mohamed, A.A.R.; Best, R.J.; Liu, X.; Morrow, D.J. Two-phase bess optimization methodology to enhance the distribution network power quality and mitigate violations. *IET Renew. Power Gener.* **2023**, *17*, 2895–2908. [\[CrossRef\]](#)
2. Zhou, F.; Lu, H.; Jiang, C. Violation with concerns of safety: A study on non-compliant behavior and the antecedent and consequent effects in power grid construction. *Saf. Sci.* **2024**, *170*, 106353. [\[CrossRef\]](#)
3. Amiri, E.; Sadeghi, S.H.H.; Moini, R. A probabilistic approach for human safety evaluation of grounding grids in the transient regime. *IEEE Trans. Power Deliv.* **2012**, *27*, 945–952. [\[CrossRef\]](#)
4. Liu, Y. Analysis of brazilian blackout on march 21st, 2018 and revelations to security for hunan grid. In Proceedings of the 2019 4th International Conference on Intelligent Green Building and Smart Grid (IGBSG), Yichang, China, 6–9 September 2019; IEEE: Beijing, China, 2019; pp. 1–5.
5. Shu, Y.; Tang, Y. Analysis and recommendations for the adaptability of china's power system security and stability relevant standards. *CSEE J. Power Energy Syst.* **2017**, *3*, 334–339. [\[CrossRef\]](#)
6. Stanescu, D.; Digulescu, A.; Ioana, C.; Serbanescu, A. On the existing and new potential methods for partial discharge source monitoring in electrical power grids. In *Smart Trends in Computing and Communications: Proceedings of SmartCom 2021*; Springer: Singapore, 2021; pp. 155–166.
7. Yang, J.; Wang, D.; Xu, J.; Li, P.; Mi, C. Research on the risk early warning of power production safety accident based on data drive. *Power Systems and Big Data* **2019**, *4*, 9–14.
8. Hu, B.; Zou, H.; Guo, R. Research on key technology of big data security analysis platform based on electric power network monitoring data. *Power Systems and Big Data* **2021**, *3*, 51–58.
9. Harimurugan, D.; Puneekar, G.S. Electric field reduction in an ehv substation for occupational exposure via transposition of conductors. *IEEE Trans. Power Deliv.* **2018**, *33*, 3147–3154. [\[CrossRef\]](#)
10. Wu, J.; Liu, M.; Zheng, Y.; Jing, S.; Huang, Q. Research on automatic generation technology for secondary equipment of security measures of smart substation. In Proceedings of the 2018 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE), Kajang, Malaysia, 29 to 31 May 2018; IEEE: Kota Kinabalu, Malaysia, 2018; pp. 223–227.
11. Long, X.; Dong, M.; Xu, W.; Li, Y.W. Online monitoring of substation grounding grid conditions using touch and step voltage sensors. *IEEE Trans. Smart Grid* **2012**, *3*, 761–769. [\[CrossRef\]](#)
12. Wu, Z.; Li, H.; Zheng, Y.; Xiong, C.; Jiang, Y.-G.; Davis, L.S. A coarse-to-fine framework for resource efficient video recognition. *Int. J. Comput. Vis.* **2021**, *129*, 2965–2977. [\[CrossRef\]](#)
13. Xu, J.; Li, K.; Li, D. An automated few-shot learning for time series forecasting in smart grid under data scarcity. *IEEE Trans. Artif. Intell.* **2024**, *5*, 2691–4581. [\[CrossRef\]](#)
14. Tsoumplekas, G.; Athanasiadis, C.L.; Doukas, D.I.; Chrysopoulos, A.; Mitkas, P.A. Few-Shot Load Forecasting Under Data Scarcity in Smart Grids: A Meta-Learning Approach. *arXiv* **2024**, arXiv:2406.05887.
15. Qian, K.; Zhao, W.; Li, K.; Ma, X.; Yu, H. Visual slam with boplv pairs using egocentric stereo camera for wearable-assisted substation inspection. *IEEE Sens. J.* **2019**, *20*, 1630–1641. [\[CrossRef\]](#)
16. Li, B.; Yang, J.; Zeng, X.; Yue, H.; Xiang, W. Automatic gauge detection via geometric fitting for safety inspection. *IEEE Access* **2019**, *7*, 87042–87048. [\[CrossRef\]](#)
17. Cai, W.; Le, J.; Jin, C.; Liu, K. Real-time image-identification-based anti-manmade misoperation system for substations. *IEEE Trans. Power Deliv.* **2012**, *27*, 1748–1754. [\[CrossRef\]](#)
18. Long, X.; Cui, W.; Zheng, Z. Safety helmet wearing detection based on deep learning. In Proceedings of the 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), Chengdu, China, 15–17 March 2019; IEEE: Beijing, China, 2019; pp. 2495–2499.
19. Liu, Y.; Liu, J.; Ke, Y. A detection and recognition system of pointer meters in substations based on computer vision. *Measurement* **2020**, *152*, 107333. [\[CrossRef\]](#)
20. Wang, H.; Zhang, X.; Sun, Y.; Li, J.; Li, Y. Research and application of artificial technology for substation environment surveillance system. In Proceedings of the 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Chongqing, China, 24–26 May 2019; IEEE: Beijing, China, 2019; pp. 901–905.

21. Patel, M.; Wang, X.; Mao, S. Data augmentation with conditional gan for automatic modulation classification. In Proceedings of the 2nd ACM Workshop on Wireless Security and Machine Learning, Linz, Austria, 13 July 2020; pp. 31–36.
22. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A.A. Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65. [[CrossRef](#)]
23. Arjovsky, M.; Bottou, L. Towards principled methods for training generative adversarial networks. *arXiv* **2017**, arXiv:1701.04862.
24. Goodfellow, I. Nips 2016 tutorial: Generative adversarial networks. *arXiv* **2016**, arXiv:1701.00160.
25. Finn, C.; Abbeel, P.; Levine, S. Model-agnostic meta-learning for fast adaptation of deep networks. In *International Conference on Machine Learning*; PMLR: Cambridge, MA, USA, 2017; pp. 1126–1135.
26. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.